心理科学进展 2018, Vol. 26, No. 2, 319–330
Advances in Psychological Science DOI: 10.3724/SP.J.1042.2018.00319

最后通牒博弈中的公平偏好：基于双系统理论的视角*

张 慧[1] 马红宇[1] 徐富明[2] 刘燕君[3] 史燕伟[1]

([1]华中师范大学心理学院，武汉 430079) ([2]江西师范大学心理学院，南昌 330022)
([3]中国人民大学劳动人事学院，北京 100872)

摘 要 在最后通牒博弈中，自利动机与公平偏好的权衡决定着个体拒绝或者接受不公平分配。根据双系统理论，对于个体拒绝不公平分配是系统 1 中的自动反应还是系统 2 中的控制加工这一争议，从理论假说、影响因素和脑机制三个方面进行了探讨。自动消极互惠假说和社会启发假说支持公平偏好的自动加工假说，公平偏好控制假说则认为公平偏好是抑制自利动机的控制加工的结果。系统 1 通过脑岛、杏仁核和腹内侧前额叶皮层识别和评估公平行为；系统 2 通过扣带回皮层、腹外侧、背内侧和背外侧前额叶皮层，对系统 1 进行重新评估和调节并做出相应的决策。个体差异因素和实验任务特征可能会影响个体在系统 1 中的自动反应。未来研究需要进一步完善实验范式，深入探索其中的调节变量及其神经网络。

关键词 公平；最后通牒博弈；双系统理论；公平偏好自动假说；公平偏好控制假说

分类号 B849:C91

1 引言

从古至今，人们都厌恶不公平对待，且愿意消耗资源来惩罚实施不公平的个体，这种追求公平的社会偏好已经受到了众多研究者的关注(Fehr & Gächter, 2000)。传统经济学模型认为个体是理性和自私的，追求个人效用最大化(Edwards, 1954)。然而，很多研究都发现个体并非是完全追求利益最大化的“经济人”，而是也会表现出关心他人收益，惩罚不公平的“社会人”，这一结果在最后通牒博弈(ultimatum game, UG)范式的实验中得到了充分的体现(Güth, Schmittberger, & Schwarze, 1982)。在 UG 中，两人共同分配一定数额的金钱，首先由提议者提出分配方案，如果接收者接受该方案，则二者按这一分配方案进行分配，如果接收者拒绝，则二者都不能获得金钱(Güth et al., 1982)。在理性框架下，根据“经济人”假设，提议者会为了自我利益提供最少的分配，接收者将接受任何非零的分配。然而，很多研究结果都与理性框架相矛盾，通常提议者会分配约 40 ~ 50%的钱，而接收者会拒绝低于 20%的分配(Camerer, 2003)。因此，研究者认为 UG 在本质上是有关公平的决策(陈叶烽，2014；王赟，吴斌，李纾，周媛，2015)，它打破了“经济人”假设中有关人类始终追求自我利益最大化的“自利”假说。

在目前研究中，很多研究者都认为在 UG 中拒绝不公平分配的行为是个体为了追求公平，放弃自我利益的社会偏好行为(Knoch & Nash, 2015; Achtziger, Alós-Ferrer, & Wagner, 2016；徐富明，李欧，邓颖，刘程浩，& 史燕伟，2016)，并分别从认知、情绪和动机等不同角度提出了相应的解释机制(罗艺，封春亮，古若雷，吴婷婷，罗跃嘉，2013)。虽然这些理论都取得了相应的实证支持，但个体拒绝不公平分配的行为是个体作为“社会人”的自动反应还是作为“经济人”为了长远利益，控制自利偏好的策略行为还存在很多争议。对这一争议的回答就需要回到一个心理学的经典争论：公平偏好行为是个体在系统 1 中的自动反应还是系统 2 中控制加工的结果。为此，研究者根据双系统理论提出了以下两种不同的研究假设：一些研究者认为，在利益冲突情境中，个体直觉、

收稿日期：2017-06-09
* 国家自然科学基金青年项目(31200795)；华中师范大学中央高校基本科研业务费项目(CCNU14Z02015)；国家自然科学基金面上项目(71571083)；2017 年华中师范大学优秀博士学位论文培育计划(2017YBZZ089)资助。
通信作者：马红宇, E-mail: mahy@mail.ccnu.edu.cn

自动的反应是公平偏好，不需要控制自利冲动，因此，拒绝不公平是个体的自动反应(Rand, Greene, & Nowak, 2012; Rubinstein, 2007)。相反，一些研究者认为个体直觉、自动的反应是自我利益，个体需要消耗认知资源、付出努力才能抵制自利冲动，因此，拒绝不公平是个体自我控制的结果(Martinsson, Myrseth, & Wollbrant, 2014; Myrseth, Fishbach, & Trope, 2009)。目前，很多研究者都从双系统理论的视角对这一问题进行了探索。

双系统理论认为个体运用两种系统对信息进行加工，在系统 1 中进行自动加工，即由刺激引发的无意识快速加工过程，是一种启发式加工，不受认知资源限制。而在系统 2 中进行控制加工，即受意识控制，需要注意的积极参与，占用系统加工资源(Evans, 2003)。和自动加工相比，控制加工更为主动和灵活，它可以随客观情况的变化不断调整资源分配(Lieberman, 2007)。双系统理论的提出与兴起引起了社会认知、判断与决策等领域众多研究者的关注，但也有研究者认为双系统的划分方式过于简单(Evans, 2008; 肖前国，罗乐，余林，2009)，不同研究者对于双系统的定义、特点、加工机制以及在具体任务中的区别等方面的理解存在着一些分歧，例如 Kruglanski 和 Gigerenzer (2011)认为双系统理论的直觉加工和审慎加工都是基于规则的。但是研究者也对双系统理论的划分提出了一些新的观点以完善和推进双系统理论的发展。例如，Evans (2008)从意识性、进化、功能性特点和个体差异四个方面对双系统的内容进行了完善。由此可见，双系统理论并非一家之言，它集众多研究者的观点于一身，并且还在不断发展完善中(罗俊龙等，2013)。在 UG 中，个体拒绝不公平的行为是个体在系统 1 中的自动反应还是系统 2 中控制加工的结果？是利益冲突情境中的自动激活反应还是深思熟虑之后的结果？

从双系统理论的角度探讨这一问题具有重要的意义。第一，相比于其他研究者分别从认知、情绪和动机等不同角度进行解释(罗艺等，2013)，双系统理论认为在 UG 中个体拒绝不公平的行为不是简单地受情绪因素所驱动，而是认知和情绪交互作用的复杂过程。双系统理论融合了社会决策中的认知因素和情绪因素，从认知和情绪交互作用的角度探讨了 UG 中个体拒绝不公平行为的原因，为解释个体拒绝不公平行为提供了一个新视角。第二，对于个体行为，研究者逐渐认识到人们的心理与行为过程不仅是个体有意识地、系统地信息加工，以建构与解释外在世界、执行行为的过程，也是受环境诱发的或无意识的自动化加工过程(Baumeister & Bargh, 2014)。双系统理论区分了意识和无意识这两种心理与行为过程(Chaiken & Trope, 1999)，从自动加工和控制加工过程来区分博弈范式下的个体公平偏好行为，有助于从根本上揭示不同研究者对个体自动反应的基本假设，拓展和丰富“经济人”“社会人”假说，也为 UG 中拒绝不公平行为的理性非理性之争提供新的证据。总之，本文旨在透过双系统理论的视角，对前人研究中关于个体在 UG 中拒绝不公平行为是自动加工还是控制加工的争议进行探讨，分别梳理 UG 中公平偏好的自动加工过程和控制加工过程及其神经机制，深入分析现有实验结果的不一致及其背后可能的原因，并提出对未来研究的展望。

2 拒绝不公平的加工过程比较

近年来，研究者已经开始探讨信息加工过程如何影响公平偏好行为(Halali, Bereby-Meyer, & Meiran, 2014; Bear & Rand, 2016)。然而，研究关于公平偏好行为是自动化的结果还是自我控制的结果尚且存在分歧。根据双系统理论，部分研究者认为公平偏好行为是系统 1 中的自动反应，然而也有一些研究者认为这一行为是受到系统 2 控制加工的表现，基于此，不同研究者提出了不同的理论假设。

2.1 自动消极互惠假说(automatic negative reciprocity hypothesis)

从广义上来说，互惠(reciprocity)是一种对他人行为进行类似回应的倾向，例如对他人好的行为给予相应的好的行为回应，即积极互惠(positive reciprocity)，或者对他人不好的行为给予不好的行为回应，即消极互惠(negative reciprocity) (Guala, 2012)。Halali 等(2014)提出的自动消极互惠假说认为个体拒绝不公平的行为是一种基于负性情绪的自动消极互惠反应。在 UG 中，面对他人的不公平分配，个体会产生负性情绪，在负性情绪下，个体可能出于消极互惠的原则拒绝他人的不公平分配。例如，van't Wout, Kahn, Sanfey 和 Aleman (2006)通过多导生理记录仪(SCR)记录个

体的情绪激活，研究发现相比于接受不公平，个体在拒绝不公平时的情绪激活更强，而且拒绝不公平与个体自我报告的愤怒情绪也呈正相关(Srivastava, Espinoza, & Fedorikhin, 2009)。此外，研究还发现个体接受不公平与情绪调节系统(Xiao & Houser, 2005)有关，情绪调节策略可降低不公平拒绝率(van't Wout, Chang, & Sanfey, 2010)。因此，个体拒绝不公平行为可能是不公平分配引发的负性情绪下的自动反应。自动消极互惠假说从情绪的角度说明了个体拒绝不公平分配行为是系统 1 中的自动反应。

2.2 社会启发假说(social heuristics hypothesis)

社会启发假说认为系统 1 的自动决策依赖于公平偏好，而系统 2 更在乎促进利益的最大化(Bear & Rand, 2016)。个体拒绝不公平是在系统 1 作用下的自动反应，但当系统 2 中的自利动机能抑制社会偏好时，个体将表现出更多的自利行为，接受不公平分配。由于系统 1 的自动加工不受认知资源的影响，而系统 2 的控制加工需要认知资源，那么在资源损耗条件下系统 2 的作用将减弱，如果此时不公平拒绝率提高，则说明个体拒绝不公平是在系统 1 作用下的自动直觉反应。因此，研究中通常运用实验操纵来减少个体的控制资源来探索这一问题。

目前研究中通常采用的方式有：第一，认知负荷。在研究中被试需要参与容易或困难的认知任务(例如，记忆一个三位数或一个七位数)，与此同时进行博弈任务。与参与简单任务的被试相比，那些参与困难任务的被试的认知负荷更大，在博弈任务中投入的认知资源更少，因此，他们的决策更多受到系统 1 中自动加工的影响，更少采用系统 2 中审慎加工的方式(Gilbert, Pelham, & Krull, 1988)。第二，时间限制。研究中通常要求被试快速做出决策(时间压力)，或决策前进行一定时间的思考(时间延迟)。时间压力减少了个体进行审慎控制加工的机会，迫使他们进行自动加工。因此，在时间压力下个体决策更多受到自动加工的影响，时间延迟时更多受到审慎控制加工的影响(Wright, 1974)。第三，自我损耗。研究中被试首先要完成消耗认知资源的任务，然后参与博弈任务，这一顺序与认知负荷的研究不同。与完成需要较少认知资源任务的被试相比，那些完成需要更多认知资源任务的被试更可能出现自我损耗，更不能抑制自动反应(Baumeister, Bratslavsky, Muraven, & Tice, 1998)。

目前，很多研究都发现，通过实验操纵使个体更多采用系统 1 中自动加工方式时，不公平拒绝率更高。例如，研究发现与决策时间充足相比，当迫使人们快速决策时，个体对不公平的拒绝率更高(Sutter, Kocher, & Strauß, 2003; Cappelletti, Güth & Ploner, 2011)，而延迟决策将降低不公平拒绝率(Grimm & Mengel, 2010; Neo, Yu, Weber, & Gonzalez, 2013)。此外，研究也发现相比于控制组，个体在睡眠剥夺(Calvillo & Burgeno, 2015)，自我控制损耗(Halali et al., 2014)以及酗酒状态下(Morewedge, Krishnamurti, & Dan, 2014)对不公平分配的拒绝率更高。此外，刘长江等(2016)也从无意识、认知能力、自我损耗等角度说明了合作行为是一种利益冲突情境中自动激活反应。由此可以说明，在资源损耗时，个体更多使用系统 1 中自动加工，自利动机减少，公平偏好行为增加，验证了公平偏好自动假说。

2.3 公平偏好控制假说(controlled-processing hypothesis)

公平偏好控制假说认为个体决策的自利动机需要被控制，面对不公平个体自动反应是自利，当系统 2 中自我控制资源能抑制自利冲动时，个体将表现出公平偏好。首先，相关研究表明自我控制能力和情绪调节能力强的个体更可能拒绝不公平(Sütterlin, Herbert, Schmitt, Kübler, & Vögele, 2011; Dunn, Evans, Makarova, White & Clark, 2012)。因为自我控制能力和情绪调节能力作为一种个体资源，使个体资源更加充足，个体将更多运用系统 2 的控制加工。并且刘长江等(2016)认为在认知资源和时间都充足的条件下，个体可能会基于理性的思考过程做出决策。即在系统 2 的控制加工下，个体能更好抑制自利动机，拒绝不公平分配。其次，Knoch, Gianotti, Baumgartner 和 Fehr (2010)的研究显示前额皮层区域的基线活动与接受不公平负相关。随着这些区域基线活动的降低，个体更少使用系统 2 的控制加工方式(Miller & Cohen, 2001)，这一结果说明个体的自动直觉反应是自利的接受不公平。此外，研究者还进一步说明了二者之间的因果关系，实验通过应用经颅磁刺激(rTMS)扰乱前额叶皮层，导致不公平分配的拒绝率降低(Knoch, Pascual-Leone, Meyer, Treyer, &

Fehr, 2006)。后续研究发现自我调节资源被消耗(Achtziger et al., 2016)和快速决策(Hochman, Ayal, & Ariely, 2015)都会促进个体使用系统 1，导致不公平拒绝率降低。由此可以说明，个体在系统 2 的控制加工条件下，更多拒绝不公平，进一步验证了公平偏好控制假说。

以上研究基于不同的理论假设都对 UG 中的公平偏好行为进行了解释，但三者之间也存在一些联系与区别(具体见表 1)：首先，在自动反应假设和支持理论观点方面，自动消极互惠假说和社会启发假说都认为公平偏好是系统 1 的自动反应，即公平偏好自动假说；而控制假说认为自利是系统 1 的自动反应，公平偏好是受控制的。其次，在双系统的作用方面，自动消极互惠假说主要从系统 1 的功能角度来进行解释，社会启发假说和控制假说都认为在公平博弈中存在自利动机和公平偏好之间的冲突，系统 2 都要在这种冲突中进行调节控制，只是二者控制的动机不同。总体来说，自利动机与公平动机的权衡决定着最终的决策行为(Civai, 2013)，虽然不同理论假设认为个体在系统 1 中的自动反应不同，但系统 2 都可以对其自动反应进行调节，当自利动机占优时，个体将更多接受不公平，当公平偏好占优时，个体将更多拒绝不公平。从目前的研究结果来看，以上理论假说都有相应的实证支持，但研究结果出现不一致的原因可能是不同研究基于不同的研究假设。目前研究中主要的假设是如果公平偏好是受控制的，那么在资源损耗条件下，系统 2 的控制功能减少，个体将更加自利，提议者将减少分配，接收者将降低不公平拒绝率。相反，如果公平偏好是自动的，在资源损耗条件下，系统 2 控制自利动机的功能减少，个体将增加公平偏好行为，提议者将增加其分配，接收者将增加不公平拒绝率。也就是说，对于系统 1 的哪种动机(自利或公平偏好)是个体的自动反应这一基本假设还存在一些争议，而个体的自动反应可能会受到很多因素的影响，下面将进一步阐述。

3 影响自动反应的因素

个体的自动反应可能是受到一些个体差异和实验任务特征因素的影响，例如社会价值取向，认知反思能力，实验轮次和实验材料等。

3.1 个体差异因素

3.1.1 社会价值取向

社会价值取向(social value orientation, SVO)是指个体在与他人进行资源分配过程中，个体对自我结果和他人结果表现出的稳定偏好(Beggan, Messick, & Allison, 1988; Murphy,

Ackermann, & Handgraaf, 2011)，主要包括以下两种类型：亲社会取向和亲自我取向。亲社会取向个体更愿意牺牲自我资源去建立公平分配和互利最大化，而亲自我取向个体更加关注自我利益，试图将自我利益最大化，忽视对他人的影响(van Lange, 1999)。前人研究发现亲社会取向个体在社会困境中更可能与他人合作(Balliet, Parks, & Joireman, 2009; Bogaert, Boone, & Declerck, 2008; van Lange, Joireman, Parks, & van Dijk, 2013)，刘长江和郝芳(2011)的研究也发现，在在初始资源处于相对劣势位置时，亲社会者比亲自我者表现出更高的合作水平。在 UG 中，Bieleke, Gollwitzer, Oettingen 和 Fischbacher (2016)的研究发现相比于系统 2 中的审慎控制加工，亲社会取向个体在自动加工方

表 **1** 最后通牒博弈中双系统理论作用的总结

理论	基本内容	双系统的作用	基本假设	支持结论	实证支持
自动消极互惠假说	个体拒绝不公平的行为可能是由于不公平分配引发的负性情绪下的自动反应。	系统 1 的功能角度来进行解释。	公平偏好是个体系统 1 中的自动反应。	公平偏好自动假说	eg. Halali et al., 2014
社会启发假说	个体拒绝不公平是个体在系统 1 作用下的自动反应，当系统 2 中的自利动机能抑制个体的社会偏好时，个体将表现出更多的自利行为。	公平博弈中存在自利动机和公平偏好之间的冲突，系统 2 能调节和控制这一冲突。	公平偏好是个体系统 1 的自动反应，但系统 2 可以影响个体的公平偏好行为。		eg. Bear & Rand, 2016
公平偏好控制假说	个体决策的自利动机是需要被控制的，面对不公平个体自动反应是自利，当系统 2 中自我控制资源能抑制自利冲动时，个体将表现出公平偏好。		自利反应是个体系统 1 的自动反应，但系统 2 可以调节和控制个体的公平偏好。	公平偏好控制假说	eg. Sütterlin et al., 2011

式下更可能拒绝不公平，说明亲社会取向个体的公平偏好更可能是一种自动反应。因此，对于亲社会取向个体来说，其自动反应是拒绝不公平来追求公平偏好，而对于亲自我取向个体来说，其自动反应可能是接受不公平追求自我利益。

3.1.2 认知反思能力

认知反思能力是指个体克服系统 1 中直觉自动加工反应，采用系统 2 中审慎控制加工的能力，一般通过用认知反思测试(Cognitive Reflection Test, CRT)来测量(Frederick, 2005; Toplak, West, & Stanovich, 2014)。在 UG 中，如果认知反思能力更高的个体更多接受不公平，说明在系统 2 的控制加工下，个体更可能接受不公平，支持了公平偏好自动假说。相反，如果认知反思能力更高的个体更多拒绝不公平，说明在系统 2 的控制加工下，个体更可能拒绝不公平，支持了公平偏好控制假说。Calvillo 和 Burgeno (2015)首先对这一问题进行了探索，结果发现认知反思测试得分更高的个体更可能接受不公平，说明在系统 2 的控制加工下，个体更加偏好自我利益，支持了公平偏好自动假说。认知反思能力与社会价值取向不同，虽然不能直接影响个体的自动反应，但不同认知反思能力的个体可能倾向于采用不同的认知加工方式，即并不是所有被试都是首先选择进行自动加工。

从个体差异的角度来看，除了个体特征因素外，一些人口学变量可能对个体的自动反应行为也会产生影响。例如，年龄(Warneken & Tomasello, 2009; Shaw & Olson, 2012)，教育背景(Carter & Irons, 1991)和受教育程度(Bellemare, Kröger, & van Soest, 2011)等因素也可能会对个体在 UG 中自动反应产生影响。

3.2 实验任务特征

3.2.1 实验轮次

在 UG 中，不同实验轮次下个体拒绝不公平行为的动机可能也不同，在重复 UG 中，拒绝不公平分配对个体来说是有利的，而在单轮 UG 中拒绝不公平分配是损失的(Rand, 2016)。因为在重复 UG 中，拒绝不公平是对他人不公平分配的一种惩罚，让他人在后续的分配中进行公平分配，从长远来看，对个体是有利的。而在单轮 UG 中，个体与他人没有后续的互动，个体拒绝不公平分配，个体收益将会减少。因此，不同实验轮次中个体拒绝不公平行为，其背后的动机可能不同。即在单轮 UG 中，个体拒绝不公平行为是为了追求公平偏好，而在重复 UG 中，个体拒绝不公平行为更可能是为了长远的利益。例如，Halali, Bereby-Meyer 和 Ockenfels (2013)的研究发现在有互惠可能的条件下，个体在自我损耗条件下有更多公平分配提议，但在无互惠可能的条件下，个体公平分配提议更少。因此，相比于单轮 UG，在多轮 UG 中，拒绝不公平行为可能不是个体的公平偏好行为，而是个体的追求自我利益的行为。虽然实验轮次不直接影响个体的自动反应，但是在不同实验轮次下，相同自动反应动机的个体，其行为决策可也能不同。例如，当个体自动反应是自我利益时，单轮 UG 中，个体可能会接受不公平，但在多轮 UG 中，个体可能会为了长远利益而拒绝不公平。目前研究主要是基于单轮 UG 的结果，在多轮 UG 中，这种动态决策过程中时间序列影响以及决策角色的变化都可能会为对现有研究结果提出新的思路，未来研究可以对这一问题进行深入探讨。

3.2.2 实验材料

目前大多数研究都是在电脑上呈现分配方案进行 UG 任务，被试可能会认为基于决策来获得金钱是小概率事件，因此，可能将分配视为假想或一个小赌注(Calvillo & Burgeno, 2015)。比起真钱，假想的分配更容易被拒绝；比起大的分配数额，小的分配更容易被拒绝，并且让被试拿着现金增加了不公平拒绝率(Shen & Takahashi, 2013)。在与公平相关的博弈中的不同操纵，可能会激发个体系统 1 中不同的动机反应，例如金钱更可能激发个体的自利动机，而公平准则启动更可能激发公平偏好的自动反应，未来研究还需从情境激发的角度探索影响系统 1 中不同动机的自动反应的因素。

4 拒绝不公平的神经机制对比

随着认知神经科学的发展，研究者开始逐步以双系统理论为基础探索其神经机制，将 UG 中个体决策反应的脑区分离成两个系统：系统 1 通过脑岛(anterior insula, AI)、杏仁核(amygdala)和腹内侧前额叶(ventral medial prefrontal cortex, vmPFC)识别和评估公平行为；系统 2 通过扣带回皮层(dorsal anterior cingulate cortex, dACC)、腹外侧前额叶(ventrolateral prefrontal cortex, vlPFC)、

背内侧前额叶(dorsomedial prefrontal cortex, dmPFC)和背外侧前额叶(dorsolateral prefrontal cortex, dlPFC)对系统 1 进行重新评估和调节，从而做出自利或公平的决策(Feng, Luo, & Krueger, 2015; Buckholtz & Marois, 2012; Sanfey & Chang, 2008)。目前，与决策加工过程一样，在神经机制的研究中，不同研究结果也分别支持了以上理论假设。

4.1 自动加工的神经机制

在 UG 中，目前研究通过对脑岛(AI)、杏仁核和腹内侧前额叶(vmPFC)功能的分析支持了公平偏好是依赖于情绪的、自动的反应(Sanfey & Chang, 2008; Feng et al., 2015)。具体来说，在 UG 中，不公平分配会激活与情绪控制相关的脑区，例如不公平分配引发的负性情绪会激活 AI，且其激活强度与不公平分配的拒绝率呈正相关(Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003)，因此，AI 通常被认为在拒绝不公平中发挥着重要作用(Tabibnia, Satpute, & Lieberman, 2008)。此外，前人研究还表明情感注入(Harlé & Sanfey, 2012)和情绪调节(Grecucci, Giorgetta, van't Wout, Bonini, & Sanfey, 2013)调节了不公平分配激活 AI 的反应，影响个体公平决策。Feng 等(2015)的元分析结果显示，在与系统 1 有关的脑区中，双侧 AI 的激活与自动直觉反应相一致，从侧面支持了公平偏好自动假说。但是，也有研究表明不公平诱发的负面情绪不能完全解释 AI 激活，因为当个体作为第三方时，不公平分配中 AI 的激活也很明显(Civai, 2013)。最近有研究探讨了 AI 中不同区域的作用(Zhou, Wang, Rao, Yang, & Li, 2014)，发现 AI 可分为背前侧、腹前侧和后侧，分别负责认知、情感以及感觉处理(Chang & Sanfey, 2013)，且背侧 AI 与认知规范违背相一致(Zhou et al., 2014)。因此，AI 在拒绝不公平行为中与情绪反应和认知规范判断都有关。

此外，杏仁核和 vmPFC 通常与社会决策和情感处理相关联(Gospic et al., 2014)。杏仁核反映了对违规行为的早期情绪反应，不公平分配条件下，激活杏仁核，个体对不公平分配拒绝率提高，而减少杏仁核激活将降低不公平拒绝率，表明杏仁核在公平偏好中发挥着作用。Xiang, Lohrenz 和 Montague (2013)研究发现被试对不公平分配的情绪反应与 vmPFC 激活相关, vmPFC 在情绪调节中起重要作用。同时，一项针对脑损伤病人的研究发现，vmPFC 损伤病人对不公平分配方案的拒绝率比正常人或者其他脑区受损的病人更高(Koenigs & Tranel, 2007)。这类病人在 UG 中的高拒绝率可能是情绪调节能力受损所导致的，在一定程度上说明了个体拒绝不公平是一种自动加工反应。

4.2 控制加工的神经机制

系统 2 主要进行深思熟虑的分析以重新评估系统 1，从而做出经济自利(Brüne et al., 2012)或公平的决策(Fehr & Camerer, 2007)。在 UG 中，公平偏好和自利动机之间的冲突导致扣带回皮层(dACC)激活，该区域负责监控经济利益和针对不公平分配直觉反应之间的冲突(Fehr & Camerer, 2007)。Civai, Crescentini, Rustichini 和 Rumiati (2012)研究表明对违反规范的不公平分配，只有当自利动机参与时，扣带回皮层才会被激活，且那些反应更强烈的个体更倾向于自利，也会体验到较大的冲突(Xiang et al., 2013; Zhou et al., 2014)。由此可以说明，在 UG 中，存在自利动机和公平偏好的冲突，且系统 2 会调节和控制系统 1 中的自动反应。

解决冲突的方式可能有两种：一方面，系统 1 中公平偏好自动反应被抑制，主要通过与情绪调节相关的脑区来实现(Silvers, Buhle, & Ochsner, 2013)，导致不公平拒绝率降低，支持公平偏好自动假说。具体来说，腹外侧前额叶(vlPFC)可能发挥自上而下的控制来调节脑岛激活(Tabibnia et al., 2008)，或者通过对具体情境的重新解释来调节脑岛反应(Silvers et al., 2013)。边缘扣带回通过自上而下的抑制杏仁核激活来解决情感冲突(Etkin, Egner, & Kalisch, 2011)。例如, Yu, Calder 和 Mobbs (2014)研究发现边缘扣带回抑制杏仁核对不公平分配负面情绪反应，边缘扣带回和杏仁核之间的这种耦合将减少不公平分配的拒绝率。

另一方面，动机冲突也可能通过克服自我利益动机拒绝不公平来解决，支持公平偏好控制假说(Baumgartner, Knoch, Hotz, Eisenegger, & Fehr, 2011)。右侧背外侧前额叶(dlPFC)被认为与系统 2 中的认知机制(Knoch & Fehr, 2007)和激活利他惩罚相关(Feng et al., 2015)。研究发现通过 rTMS 中断右侧 dlPFC 将会减弱公平偏好(Knoch et al., 2006)。因此，在不公平条件下被试的首要冲动是自我利益, dlPFC 的功能是抑制自利而非情绪。此

外, 前额叶和扣带回皮层除了执行认知控制的角色(系统 2), 这些脑区可能有助于整合不同信息源(例如, 预期), 并在特定的社会背景中优化反应选择(Sanfey, Stallen, & Chang, 2014)。“融合和选择”的假设也认为系统 2 是与环境相联系的, 在不同的社会环境中右侧 dlPFC 的激活中对不公平分配造成了相反的效果：在社会规范遵守的背景下将加强对不公平分配的拒绝率, 但在自愿遵守规范下将减少对不公平分配的拒绝率(Grecucci et al., 2013)。

综上, 在 UG 决策中涉及两个分离而又相互作用的脑功能区。系统 1(AI 和 vmPFC)与不公平分配唤起消极情绪引发的自动直觉反应相关; 那些与自身经济利益相冲突的直觉反应导致扣带回皮层激活, 系统 2 通过激活或抑制系统 1, 对直觉反应进行认知控制(vlPFC, dACC, dmPFC, 左侧 dlPFC)或克服自我利益(右侧 dlPFC)。目前研究存在争议的原因主要有：第一, 不同研究基于不同的前提假设。第二, 系统 2 作用不同。在双系统理论中, 系统 1 主要是支持了自动消极互惠假说, 但在系统 2 中的控制加工更加灵活, 既可以减少系统 2 的作用来降低自利动机支持社会启发假说, 也可以克服自利冲突支持控制假说。第三, 脑区功能解释不同。例如, 研究者认为背外侧前额叶激活时, 个体对不公平接受率提高, 是因为背外侧前额叶调节控制了个体的公平偏好, 支持了自动反应假说; 但也有研究者认为抑制背外侧前额叶, 个体对不公平接受率也会提高, 是因为背外侧前额叶抑制了自利动机, 支持了控制假说。那么, 对于背外侧前额叶在公平博弈中如何发挥作用还有待进一步分析论证。罗艺等(2013)认为左侧和右侧背外侧前额叶在个体执行与遵守公平准则时可能起着不同的作用, 未来的研究则需要进一步分离左侧和右侧背外侧前额叶在公平相关决策中所对应的心理机制。个体在 UG 中的决策可能是取决于系统 1 和系统 2 之间的动态平衡, 当背外侧前额叶占优势时, 系统 2 中认知过程抑制负性情绪, 从而出现自利行为; 当系统 1 占优势时, 消极情绪冲破了认知过程, 在情绪诱导下出现公平偏好。

5　研究展望

5.1　进一步完善实验范式

首先, 在与公平有关的博弈研究中包括两类不公平：一是劣势不公平, 即作为接受者, 在分配方案中的分配数额少于提议者; 一类是优势不公平, 即在分配方案中接收者的分配数额多于提议者。目前大多数研究都是基于对劣势不公平的分析, 而个体是否更可能接受优势不公平？未来研究可以同时操纵两类不公平以分析其心理机制。其次, 对于目前用来探讨不同加工方式影响个体社会决策的实验证据同样需要关注。例如, 部分研究者采用反应时间来操纵资源损耗, 但也有研究者认为反应时间反映的是决策冲突, 而不是不同的加工方式(Evans, Dillon, & Rand, 2015; Krajbich, Bartling, Hare, & Fehr, 2015), 相比于缓慢决策, 快速决策并不一定反映的是个体采用了系统 1 中的自动加工方式。但是, 在研究中可以通过不同的处理诱发个体不同的加工方式, 让个体对不公平的行为反应和反应时间不同(Lotz, 2015; Rand et al., 2012)。因此, 在时间限制的研究中可以明确告诉被试使用自动还是控制加工, 并让被试在正式实验之前进行一些练习, 减少直觉加工与理解的混淆, 并在研究中注意要在博弈决策期间采用时间压力的操纵。此外, 对于不同加工方式对 UG 中决策的影响, 也可以直接采用诱导加工范式。在研究中要求被试使用直觉自动或审慎控制的加工方式进行决策(Liu & Hao, 2011; Shenhav, Rand, & Greene, 2012), 从而更直接地探讨个体在 UG 中拒绝不公平行为是系统 1 还是系统 2 的作用。最后, 在公平偏好控制假说中通常会采用 rTMS, 虽然生理研究是对社会决策过程研究的有利补充, 并提供了有力的实验控制(Rilling & Sanfey, 2011), 但反过来, 与其他操纵(例如, 认知负荷)相比, 生态效度更低。因此, 阻碍前额皮层激活是否可以直接与直观模式等同起来还不明确, 或者说这一操纵只反映了一种单纯的缺少系统 2 的作用。因此, 未来的研究还需要进一步完善实验范式, 得出有效一致的结论。

5.2　对双系统在公平博弈中的调节变量进行探索

双系统理论说明了加工方式会影响个体在公平博弈中的决策, 但对于系统 1 中自动反应的不一致可能是由于存在其他调节变量。首先, 从个体差异的角度来看, 不同个体可能存在不同的自动反应行为, 未来研究可以对教育程度、社会距离、人格特质等变量进行探索。因为个体行为是

否需要进行控制，可能取决于个体的行为偏好(Knoch & Nashi, 2015)，当个体的自动反应是自利时，其公平偏好行为才需要系统 2 进行控加工。其次，从个体资源的角度来看，个体拥有的资源不同可能也会对公平博弈决策产生影响，例如Tricomi, Rangel, Camerer 和 O'Doherty (2010)的研究发现个体初始财富会对公平偏好产生影响，那么对于资源拥有较少，如贫困等弱势群体而言，自动反应是否更加自利？未来研究可以尝试对拥有不同资源的个体进行对比。对于不同个体来说，个体的自我控制能力，情绪调节能力，都可能作为一种个体资源来影响个体的加工方式。例如，自我控制能力更强的个体可能更不容易出现自我损耗(范伟等, 2016)。此外，研究也发现特殊人群的某些脑区的结构或功能的异常会影响他们的社会功能，阻碍他们做出适当的社会决策(郭秀艳，郑丽，程雪梅，刘映杰，李林，2017)。因此，影响系统 1 和系统 2 中功能的因素可能不同，在系统 1 中，个体可能会受到个体或环境因素的影响，激发出不同的动机取向，在一定程度上为公平偏好自动和控制加工的争议提供了可能的解释路径。在系统 2 中更可能受到相应资源的影响，那么是否不同的资源损耗都会带来同样的结果，或者补充资源能否缓解这一结果，这些在未来研究中都需要进一步探索。

5.3 深入探索双系统理论下的神经基础

在 UG 脑机制的研究中大多是对单个脑区进行定位和解释，还需要对于脑区之间的关系和功能进行探讨，从整体上对公平博弈的神经网络结构和功能进行研究。例如，在不公平分配中，个体的情绪、认知脑区都会被激活，脑岛、背外侧前额叶、扣带回皮层这几个脑区可能存在相互作用，反映了个体决策的整合神经网络。但对于各个脑区之间的关系和冲突还需要进行探索，进一步理清各个脑区之间的关联，为双系统理论下的神经机制研究奠定基础。同时尝试结合不同脑区的功能形成神经通路系统，例如 Servaas 等(2015)认为脑岛、扣带回皮层，前额叶和右背侧纹状体构成了一个刺激−行动−奖励和动机激活系统，不公平分配将激活与社会规范相关的脑区，而决策将激活与奖励和动机相关的脑区。因此，在 UG 中，系统 1 可能在不公平分配的激活中起着重要作用，而系统 2 可能在行为决策中发挥功能。未来研究可以尝试采用不同的分析方法，如 Granger 因果模型分析法，动态因果模型等方法来探讨不同脑区之间的协同性或同步性，以及阐明脑区之间信息的方向性传递。

总之，对于最后通牒博弈中拒绝不公平行为是个体的自动反应还是个体控制加工后的决策这一争议，基于双系统理论的视角，不同研究者都对这一问题进行了探讨。虽然目前研究并未得出一致结论，但研究者都认为在 UG 中，个体存在自利动机和公平偏好的冲突，且系统 2 会调节和控制系统 1 中的自动反应。

致谢：感谢中国人民大学劳动人事学院博士生张柏楠对英文摘要的修改，也感谢审稿专家的修改意见！

参考文献

陈叶烽. (2014). 最后通牒实验与人类的公平感. *南方经济*, (6), 81−86.

范伟, 钟毅平, 李慧云, 孟楚熠, 游畅, 傅小兰. (2016). 欺骗判断与欺骗行为中自我控制的影响. *心理学报, 48*(7), 845−856.

刘长江, 郝芳. (2011). 不对称社会困境中社会价值取向对合作的影响. *心理学报, 43*(4), 432−441.

刘长江, 张跃, 郝芳, 刘采梦, 丁絮, 石雨. (2016). 利益冲突情境中社会行为的自动激活: 合作还是利己?. *心理科学进展, 24*(12), 1897−1906.

罗俊龙, 张恩涛, 岳彩镇, 唐晓晨, 钟俊, 张庆林. (2013). 基于双加工理论解释下信念偏差效应的神经机制. *心理科学进展, 21*(5), 800−807.

罗艺, 封春亮, 古若雷, 吴婷婷, 罗跃嘉. (2013). 社会决策中的公平准则及其神经机制. *心理科学进展, 21*(2), 300−308.

郭秀艳, 郑丽, 程雪梅, 刘映杰, 李林. (2017). 不公平感及相关决策的认知神经机制. *心理科学进展, 25*(6), 903−911.

王赟, 吴斌, 李纾, 周媛. (2015). 公平的神经基础. *科技导报, 33*(9), 83−92.

肖前国, 罗乐, 余林. (2009). 推理与决策的双加工理论研究简评. *心理科学进展, 17*(2), 321−324.

徐富明, 李欧, 邓颖, 刘程浩, 史燕伟. (2016). 行为经济学中的不平等规避. *心理科学进展, 24*(10), 1613−1622.

Achtziger, A., Alós-Ferrer, C., & Wagner, A. K. (2016). The impact of self-control depletion on social preferences in the ultimatum game. *Journal of Economic Psychology, 53*, 1−16.

Balliet, D. P., Parks, C., & Joireman, J. J. (2009). Social value

orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations, 12*(4), 533−547.

Baumeister, R. F., & Bargh, J. A. (2014). Conscious and unconscious: Toward an integrative understanding of human mental life and action. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 35–49). New York, NY: Guilford Press.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource?. *Journal of Personality and Social Psychology, 74*(5), 1252−1266.

Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., & Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience*, *14*(11), 1468−1474.

Bear, A., & Rand, D. G. (2016). Intuition, deliberation, and the evolution of cooperation. *Proceedings of the National Academy of Sciences, 113*(4), 936−941.

Beggan, J. K., Messick, D. M., & Allison, S. T. (1988). Social values and egocentric bias: Two tests of the might over morality hypothesis. *Journal of Personality and Social Psychology, 55*(4), 606−611.

Bellemare, C., Kröger, S., & van Soest, A. (2011). Preferences, intentions, and expectation violations: A large-scale experiment with a representative subject pool. *Journal of Economic Behavior & Organization, 78*(3), 349−365.

Bieleke, M., Gollwitzer, P. M., Oettingen, G., & Fischbacher, U. (2016). Social value orientation moderates the effects of intuition versus reflection on responses to unfair ultimatum offers. *Journal of Behavioral Decision Making,* in press.

Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology, 47*(3), 453–480.

Brüne, M., Scheele, D., Heinisch, C., Tas, C., Wischniewski, J., & Güntürkün, O. (2012). Empathy moderates the effect of repetitive transcranial magnetic stimulation of the right dorsolateral prefrontal cortex on costly punishment. *PLoS One, 7*(9), e44747.

Buckholtz, J. W., & Marois, R. (2012). The roots of modern justice: Cognitive and neural foundations of social norms and their enforcement. *Nature Neuroscience, 15*(5), 655−661.

Calvillo, D. P., & Burgeno, J. N. (2015). Cognitive reflection predicts the acceptance of unfair ultimatum game offers. *Judgment and Decision Making, 10*(4), 332−341.

Camerer, C. F. (2003). Strategizing in the brain. *Science, 300*(5626), 1673−1675.

Cappelletti, D., Güth, W., & Ploner, M. (2011). Being of two minds: Ultimatum offers under cognitive constraints. *Journal of Economic Psychology, 32*(6), 940−950.

Carter, J. R., & Irons, M. D. (1991). Are economists different, and if so, why? *Journal of Economic Perspectives, 5*(2), 171–177.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Chang, L. J., & Sanfey, A. G. (2013). Great expectations: Neural computations underlying the use of social norms in decision-making. *Social Cognitive and Affective Neuroscience, 8*(3), 277−284.

Civai, C. (2013). Rejecting unfairness: Emotion-driven reaction or cognitive heuristic?. *Frontiers in Human Neuroscience, 7*, 126.

Civai, C., Crescentini, C., Rustichini, A., & Rumiati, R. I. (2012). Equality versus self-interest in the brain: Differential roles of anterior insula and medial prefrontal cortex. *NeuroImage, 62*(1), 102−112.

Dunn, B. D., Evans, D., Makarova, D., White, J., & Clark, L. (2012). Gut feelings and the reaction to perceived inequity: The interplay between bodily responses, regulation, and perception shapes the rejection of unfair offers on the ultimatum game. *Cognitive, Affective, & Behavioral Neuroscience, 12*(3), 419−429.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*(4), 380−417.

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences, 15*(2), 85−93.

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454–459.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255−278.

Evans, A. M., Dillon, K. D., & Rand, D. G. (2015). Fast but not intuitive, slow but not reflective: Decision conflict drives reaction times in social dilemmas. *Journal of Experimental Psychology: General, 144*(5), 951−966.

Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: The neural circuitry of social preferences. *Trends in Cognitive Sciences, 11*(10), 419−427.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives, 14*(3), 159−182.

Feng, C., Luo, Y.-J., & Krueger, F. (2015). Neural signatures of fairness-related normative decision making in the

ultimatum game: A coordinate-based meta-analysis. *Human Brain Mapping, 36*(2), 591−602.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25−42.

Gilbert, D. T., Krull, D. S., & Pelham, B. W. (1988). Of thoughts unspoken: Social inference and the self-regulation of behavior. *Journal of Personality and Social Psychology, 55*(5), 685−694.

Gospic, K., Sundberg, M., Maeder, J., Fransson, P., Petrovic, P., Isacsson, G., … Ingvar, M. (2014). Altruism costs—the cheap signal from amygdala. *Social Cognitive and Affective Neuroscience, 9*(9), 1325−1332.

Grecucci, A., Giorgetta, C., van't Wout, M., Bonini, N., & Sanfey, A. G. (2013). Reappraising the ultimatum: An fMRI study of emotion regulation and decision making. *Cerebral Cortex*, *23*(2), 399−410.

Grimm, V., & Mengel, F. (2010). Let me sleep on it: Delay reduces rejection rates in ultimatum games. *Economics Letters, 111*(2), 113−115.

Guala, F. (2012). Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences, 35*(1), 1−15.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3*(4), 367−388.

Halali, E., Bereby-Meyer, Y., & Meiran, N. (2014). Between self-interest and reciprocity: The social bright side of self-control failure. *Journal of Experimental Psychology: General, 143*(2), 745−755.

Halali, E., Bereby-Meyer, Y., & Ockenfels, A. (2013). Is it all about the self? The effect of self-control depletion on ultimatum game proposers. *Frontiers in Human Neuroscience, 7*, 240.

Harlé, K. M., & Sanfey, A. G. (2012). Social economic decision-making across the lifespan: An fMRI investigation. *Neuropsychologia, 50*(7), 1416−1424.

Hochman, G., Ayal, S., & Ariely, D. (2015). Fairness requires deliberation: The primacy of economic over social considerations. *Frontiers in Psychology, 6*, 747−753.

Knoch, D., & Fehr, E. (2007). Resisting the power of temptations. *Annals of the New York Academy of Sciences, 1104*, 123−134.

Knoch, D., Gianotti, L. R. R., Baumgartner, T., & Fehr, E. (2010). A neural marker of costly punishment behavior. *Psychological Science, 21*(3), 337−342.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science, 314*(5800), 829−832.

Knoch, D., & Nash, K. (2015). Self-control in social decision making: A neurobiological perspective. In G. Gendolla, M. Tops, & S. Koole (Eds.), *Handbook of Biobehavioral Approaches to Self-Regulation* (pp. 221-234). New York: Springer.

Koenigs, M., & Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the ultimatum game. *Journal of Neuroscience, 27*(4), 951–956.

Krajbich, I., Bartling, B., Hare, T., & Fehr, E. (2015). Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications, 6*, 7455.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review, 118*(1), 97−109.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology, 58*, 259−289.

Liu, C.-J., & Hao, F. (2011). An application of a dual-process approach to decision making in social dilemmas. *The American Journal of Psychology, 124*(2), 203−212.

Lotz, S. (2015). Spontaneous giving under structural inequality: Intuition promotes cooperation in asymmetric social dilemmas. *PLoS One, 10*(7), e0131562.

Martinsson, P., Myrseth, K. O. R., & Wollbrant, C. (2014). Social dilemmas: When self-control benefits cooperation. *Journal of Economic Psychology, 45*, 213−236.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167−202.

Morewedge, C. K., Krishnamurti, T., & Dan, A. (2014). Focused on fairness: Alcohol intoxication increases the costly rejection of inequitable rewards. *Journal of Experimental Social Psychology, 50*, 15−20.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making, 6*(8), 771−781.

Myrseth, K. O. R., Fishbach, A., & Trope, Y. (2009). Counteractive self-control: When making temptation available makes temptation less tempting. *Psychological Science, 20*(2), 159−163.

Neo, W. S., Yu, M., Weber, R. A., & Gonzalez, C. (2013). The effects of time delay in reciprocity games. *Journal of Economic Psychology, 34*, 20−35.

Rand, D. G. (2016). Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science, 27*(9), 1192−1206.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature, 489*(7416), 427−430.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of

social decision-making. *Annual Review of Psychology, 62*(1), 23−48.

Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *The Economic Journal, 117*(523), 1243−1259.

Sanfey, A. G., & Chang, L. J. (2008). Multiple systems in decision making. *Annals of the New York Academy of Sciences, 1128*, 53−62.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755−1758.

Sanfey, A. G., Stallen, M., & Chang, L. J. (2014). Norms and expectations in social decision-making. *Trends in Cognitive Sciences, 18*(4), 172−174.

Servaas, M. N., Geerligs, L., Renken, R. J., Marsman, J.-B., Ormel, J., & Riese, H., et al. (2015). Connectomics and neuroticism: An altered functional network organization. *Neuropsychopharmacology, 40*(2), 296−304.

Shaw, A., & Olson, K. R. (2012). Children discard a resource to avoid inequity. *Journal of Experimental Psychology General, 141*(2), 382−395.

Shen, J., & Takahashi, H. (2013). A cash effect in ultimatum game experiments. *The Journal of Socio-Economics, 47*, 94−102.

Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in god. *Journal of Experimental Psychology General, 141*(3), 423−428.

Silvers, J. A., Buhle, J. T., & Ochsner, K. N. (2013). The neuroscience of emotion regulation: Basic mechanisms and their role in development, aging, and psychopathology. In K. N. Ochsner & K. Stephen (Eds.), *The oxford handbook of cognitive neuroscience* (pp. 53−78). New York, NY: Oxford University Press.

Srivastava, J., Espinoza, F., & Fedorikhin, A. (2009). Coupling and decoupling of unfairness and anger in ultimatum bargaining. *Journal of Behavioral Decision Making, 22*(5), 475−489.

Sutter, M., Kocher, M., & Strauß, S. (2003). Bargaining under time pressure in an experimental ultimatum game. *Economics Letters, 81*(3), 341−347.

Sütterlin, S., Herbert, C., Schmitt, M., Kübler, A., & Vögele, C. (2011). Overcoming selfishness: reciprocity, inhibition, and cardiac-autonomic control in the ultimatum game. *Frontiers in Psychology, 2*, 173.

Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science, 19*(4), 339−347.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*(2), 147−168.

Tricomi, E., Rangel, A., Camerer, C. F., & O'Doherty, J. P. (2010). Neural evidence for inequality-averse social preferences. *Nature, 463*(7284), 1089−1091.

van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology, 77*(2), 337−349.

van Lange, P. A. M., Joireman, J., Parks, C. D., & van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes, 120*(2), 125–141.

van't Wout, M., Chang, L. J., & Sanfey, A. G. (2010). The influence of emotion regulation on social interactive decision-making. *Emotion, 10*(6), 815−821.

van't Wout, M., Kahn, R. S., Sanfey, A. G., & Aleman, A. (2006). Affective state and decision-making in the ultimatum game. *Experimental Brain Research, 169*(4), 564−568.

Warneken, F., & Tomasello, A. M. (2009). Varieties of altruism in children and chimpanzees. *Trends in Cognitive Sciences, 13*(9), 397−402.

Wright, P. (1974). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology, 59*(5), 555−561.

Xiang, T., Lohrenz, T., & Montague, P. R. (2013). Computational substrates of norms and their violations during social exchange. *Journal of Neuroscience, 33*(3), 1099−1108.

Xiao, E., & Houser, D. (2005). Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences of the United States of America, 102*(20), 7398−7401.

Yu, R., Calder, A. J., & Mobbs, D. (2014). Overlapping and distinct representations of advantageous and disadvantageous inequality. *Human Brain Mapping, 35*(7), 3290−3301.

Zhou, Y., Wang, Y., Rao, L.-L., Yang, L.-Q., & Li, S. (2014). Money talks: Neural substrate of modulation of fairness by monetary incentives. *Frontiers in Behavioral Neuroscience, 8*, 150..

Fairness preferences in the Ultimate Game: A dual-system theory perspective

ZHANG Hui[1]; MA Hong-yu[1]; XU Fu-ming[2]; LIU Yanjun[3]; Shi Yan-wei[1]

([1] *School of Psychology, Central China Normal University, Wuhan* 430079, *China*)

([2] *School of Psychology, Jiangxi Normal University, Nanchang* 330022, *China*)

([3] *School of Labor and Human Resources, Renmin University of China, Beijing* 100872, *China*)

Abstract: The ultimatum game is commonly used to examine fairness-related economic decision making, and the trade-off between fairness preferences and self-interest is assumed to determine whether individuals reject or accept unfair offers. With respect to the dual-system theory, there are controversial understandings on whether fairness preferences result from the automatic response in System 1 or the deliberation processes in System2. Our study discussed such controversy from three aspects of this theory, including theoretical hypotheses, influential factors, and neural mechanisms. The automatic negative reciprocity hypothesis and the social heuristics hypothesis contend that fairness preferences are automatic, whereas the controlled-processing hypothesis contends that fairness preferences are products of deliberation process that suppresses self-interest motivation. System 1 identifies and evaluates fairness via anterior insula, amygdala, and ventromedial prefrontal cortex; while System 2 reassesses and adjusts System 1 to make the final decision via dorsal anterior cingulate cortex, ventrolateral PFC, dorsomedial PFC, and left dorsolateral PFC. Individual differences and experimental task characteristics may affect individuals' automatic responses in System 1. Future research need to further improve the experimental paradigm; explore the moderators within the dual system and its neural network.

Key words: fairness; ultimate game; dual systems theory; automatic-processing hypothesis; controlled-processing hypothesis